

Representing Medical Knowledge in a Terminological Language is Difficult¹

Ira J. Haimowitz

Ramesh S. Patil

Peter Szolovits

Clinical Decision Making Group, MIT Laboratory for Computer Science
545 Technology Square, Cambridge, MA 02139

Abstract

We report on an experiment to use a modern knowledge representation language, NIKL, to express the knowledge of a sophisticated medical reasoning program, ABEL. We are attempting to put the development of more capable medical programs on firmer representational grounds by moving from the *ad hoc* representations typical of current programs toward more principled representation languages now in use or under construction. Our experience with the project reported here suggests caution, however. Attempts at cleanliness and efficiency in the design of representation languages lead to a poverty of expressiveness that makes it difficult if not impossible to say in such languages what needs to be stated to support the application.

1. Introduction

Despite about fifteen years' efforts to build expert-level medical reasoning systems, the field has exhibited only a very small number of successful applications in very narrow fields[4]. Most projects under current development promise additional successes also in narrow, well-defined areas of medicine[14]. We believe that dramatic broad progress will come about only when we can build programs that incorporate a large variety of types of medical knowledge and that apply that knowledge in correspondingly flexible ways. Representation methods typical of many of the early medical AI programs include:

1. Diseases and physiologic states described as patterns of manifestations.
2. Rules for incrementally advancing the diagnostic process: (a) from signs and symptoms to interpreted states, (b) on to general diagnostic hypotheses, and finally (c) refining these to final diagnoses.

Each of these is typically augmented by some measure of likelihood, either of the association or the local inference. Experience with such programs suggests that they cannot generally handle the most difficult medical problems—those involving overlapping disorders and therapies[17]. Therefore, contemporary

programs often include explicit knowledge of causality, probabilistic relationships, temporal progressions, anatomy, physiology and pathophysiology, to enable them to reason more deeply about difficult, unforeseen problems. Unfortunately, the simple knowledge representations devised for the early programs were not designed to deal with the complexities introduced by these additional forms of knowledge. As a result, even the most capable of these current programs often rests on expedient incremental extensions of *ad hoc* representational methods. Therefore, exploitation of the additional power to be gained from these additional forms of knowledge is still very limited, and the need for a better means of knowledge representation appears on the critical path.

Starting with Woods' seminal critique of the uncritical use of semantic networks for representing knowledge[19], AI researchers have worked toward establishing some principles of what should count as a suitable knowledge representation language. We list here some of the desiderata that seem to be accepted by many in the knowledge representation community. Items early on our list are quite generally accepted, whereas some later points are more specific to the NIKL[3, 6, 10] language.

1. The system should have a precise semantics, either based on the semantics of first-order predicate calculus or at least defined with comparable precision. The "it means what the system does with it" point of view is explicitly rejected.
2. The knowledge representation system should automatically provide certain logical inferences. This is what sets a knowledge representation apart from a conventional data base: it can answer for the user questions beyond what was explicitly told to the system.
3. Almost universal is the notion that some form of taxonomic inheritance should be provided. This supports the inheritance of characteristics and relationships by more specialized concepts from more general ones. Possible variations here include whether the taxonomy is a pure tree or a more general graph, just what can actually be inherited, and whether inheritance admits of exceptions.
4. In order to support the logical intelligibility of the automatically provided inferences, the inference mechanism should be sound (deriving no false conclusions from true

¹The work reported here has been supported (in part) by National Institutes of Health grants R01 LM 04493 from the National Library of Medicine and R24 RR 01320 from the Division of Research Resources.

knowledge) and complete (guaranteeing that all true conclusions within the class automatically promised will in fact be made).

5. Whatever automatic inference is provided by the representation system itself should be particularly efficient, in comparison to the operation of a more general inferential machine. This suggests that the system's reasoning be partitioned among a small number of different reasoning mechanisms, one of which provides the automatic inferences. In NIKL and related systems, this motivates the distinction between the *terminological* and *assertional* reasoners. The rationale here is that one of the central concerns in building intelligent systems is the control of where effort is to be expended. Thus, a low-level part of the system such as the knowledge representation should never engage in expensive computations.
6. As a result of the demands for soundness, completeness, and efficiency, the expressive power of the representation must be so reduced that undecidable and NP-hard questions can never be raised within the representation.

In spite of these apparently restrictive desiderata, many argue that languages based on these principles are the appropriate basis for the large intelligent systems of the future[7].

Because of our serious need for a better representational basis and because of the availability of the NIKL language, which is one instance of a language that reflects the above viewpoint, we undertook to translate the existing medical knowledge base of ABEL[11, 12, 13], a program for expert consultation on disorders of acid/base and electrolyte balance, into NIKL. In the process, we hoped to explore the practicality of the knowledge representation viewpoint exemplified by NIKL, which has, after all, not been tested on any significant sized medical knowledge base. Further, if practical, we hoped to improve ABEL by cleaning up its underlying knowledge base and preparing it for further augmentations.

2. Re-representing the ABEL Knowledge Base

2.1. Synopsis of ABEL

ABEL is a knowledge-based system for diagnosis of acid-base and electrolyte disorders. Its knowledge base contains important classes of medical entities, general relationships between these classes, and formulas for calculating parameter values critical to diagnosis. ABEL accesses its knowledge base to build a diagnostic model of the patient, for formulating and evaluating hypotheses.

The knowledge base contains definitions and descriptions of electrolyte and acid-base disorders, and of other diseases which are either partial or ultimate etiologies of these disorders. There is also knowledge about pertinent body fluids, electrolytes within these fluids, and ranges of values describing low, normal, and high concentrations of these electrolytes. To model human physiology, ABEL contains descriptions of relevant anatomical organs and systems, and relationships (such as part of, connected to, and spatially inside and outside) between these components. The knowledge base tries to give a complete description of relevant human physiology without overburdening the program with in-

formation that physicians do not use in diagnosis.

Two important representations underlie the ABEL knowledge base: causal links and hypotheses at multiple levels of detail. To model causal reasoning, an important aspect of an expert physician's thinking, there are *causal links* linking diseases to acid-base and electrolyte states and to each other. ABEL uses these links to construct a causal pathway from the symptoms (abnormal acid-base states) back to the ultimate etiologies. To model reasoning at multiple levels of pathophysiological detail, also valuable to diagnosis, portions of the knowledge base are divided into five levels of pathophysiological detail. At the most detailed, *pathophysiological* level, disorders are represented as ranges of numerical values or as increases or decreases of variable values. At the least detailed, *clinical* level, disorders and their etiologies are represented by more general terms. The more pathophysiologically detailed a level is, the longer the chains of causal links between states and diseases are. *Focal links* connect selected concepts at each level to concepts at the next higher and next lower levels of detail. This multi-level structure is useful for organizing knowledge, for generating explanations at various levels of detail, and for modeling a physician's diagnostic reasoning.

ABEL's main data structure used in diagnosis is the *patient specific model* (PSM). A PSM represents a possible diagnosis for the disorders of one particular patient. The PSM is constructed by instantiating portions of the knowledge base; thus a PSM is a five-level causal network of instantiated concepts (or *nodes*), causal links, and focal links. A given patient will usually have several hypothesized diagnoses, corresponding to multiple PSMs.

To reason about multiple etiologies for the same disorders, ABEL uses *component decomposition*, in which a disorder node is divided into two constituents, one already known to be present, and one still to be confirmed. After decomposing a disorder into components, ABEL switches its diagnostic goal from resolving the original disorder to resolving that disorder's unknown component.

ABEL diagnoses by first creating initial PSM(s) for each patient, containing instances for the laboratory data. Then ABEL asks additional information of the user so that it can add more detail to the hypotheses, rank them, or confirm one of them. A PSM is scored according to how many of the original acid-base and electrolyte disorders it accounts for, and how many it fails to explain. Upon finding PSM(s) with sufficiently high score(s), ABEL reports a successful diagnosis, and explains its results to the user.

2.2. Methods of Re-representation

While re-representing the ABEL knowledge base in NIKL, we kept several important principles in mind:

The knowledge base should have a structure similar to the domain being modeled. This view has been expressed in various forms, such as Smith's Knowledge Representation Hypothesis [15, page 2]. In medical expert systems, this criterion is particularly important. In order to be accepted by the medical community, computer programs must be able to diagnose, recommend therapy, and explain their reasoning based on first principles of physiology and causality [4, page 467]. This will only be possible if the program's representation of complex medical relationships

example, the representation of a part-whole relationship should include transitivity of that relation, and that an entity may be part of multiple other entities.

There should be a separation of definitional and assertional knowledge. For example, a blood serum electrolyte has a concentration property, describing its percentage of the blood serum, but this concentration is not part of the defined meaning of a blood serum electrolyte. Knowledge related to a term's definition should be kept distinct from other information about that term, for several reasons. First, people can distinguish in their minds between essential properties of an object and other, incidental facts. Second, when the user of a system asks for a definition of a term, the program should not in turn describe everything known about it; this is inaccurate, and potentially overwhelming for the user. Third and most worrisome is that combining terminological and assertional knowledge can result in faulty classification of new terms in the taxonomies. An example is given in [5, page 40].

As many concepts and roles as possible should be completely defined. Medicine is a domain especially marked by many "primitive" terms which can not be completely defined with a set of necessary and sufficient conditions, such as **person**, **fluid** and **disease**. Nevertheless, because the power of NIKL is its use as a terminological system, we tried to define completely as many terms as possible.

The knowledge base should be modular. With the code for the knowledge base divided into smaller modules according to content, it is easier to read, understand, and revise if necessary. Thus in the ABEL knowledge base there are different portions of code for basic fluid definitions, anatomical relations, diseases, etiologies, and causal relations.

With these crucial points in mind, we re-implemented the ABEL knowledge base code in NIKL, one module at a time. The coding took approximately three person-months; minor refinements were made occasionally thereafter. The resulting knowledge base is large and richly structured; it contains approximately 1600 basic concepts and about 120 role relations interconnecting them.

3. Results

The NIKL knowledge representation was not expressive enough for the many types of knowledge in ABEL. Although some knowledge was conveniently represented, most of the information was forced inadequately into NIKL, or not represented at all.

3.1. Useful Representations

NIKL concept taxonomies were valuable for their structured hierarchies and inheritance of definitional attributes. ABEL contains deep taxonomies for diseases, electrolytes, and attributes. Role restrictions were useful for completely defining certain types of concepts, such as **low-serum-K**, defined by its parent **serum-K**, and the restriction of its **value** role to **low**. Number restrictions were good for modeling some role fillers, like the value of a parameter, as unique. NIKL disjointness classes and covers were useful in a limited way. We used disjointness classes to distinguish the different fluids and electrolytes, and we combined disjointness classes with covers to represent partitions. For ex-

ample, infectious etiologies are partitioned into **bacterial**, **rickettsial**, **viral** and **fungal**. Although covers and disjointness classes are propagated by the classifier to some related concepts, several other important inferences, including determining if a concept is incoherent, are not made.

3.2. Troublesome Representations

Most of ABEL's knowledge base, however, could not be adequately represented. Anatomical relationships were particularly troubling. Part-whole relations, both for substructure and functional element, could not be easily represented so that the important associated inferences were made. To approximate an accurate model of part-whole we used NIKL role restrictions. Thus "a nephron is part of a kidney" was represented by attaching a **nephron** role to the **kidney** concept. This representation embodied neither the transitivity nor the multi-valued properties of part-whole. Also poorly represented were containment, connection, and spatial relationships in human anatomy. These were represented with "link-concepts" having from and to roles, specifying what concepts the link connects; e.g., **thoracic-cavity**—**contains**—>**heart-space**. Such a concept is a meaningless "definition" that really represents an assertion about the two linked items. Causation, which in reality comes in a variety of patterns, could not be represented accurately for even one type. "Causal link concepts" were similar to those for containment relations, and were equally inaccurate representations. In general, when representing transitive and multi-valued relations NIKL could not automatically deduce all appropriate inferences.

Another major difficulty was that the patient-specific models (PSM's) used in ABEL's diagnostic reasoning could not be reasonably represented. The chief reason for this is NIKL's lack of support for representing instances; thus, although the concept **kidney-disease** might be present, there is no place for the instance representing a particular kidney disease in a particular patient. More fundamental were difficulties in representing quantitative relations such as component decomposition—e.g., that the total potassium deficit from two concurrent causes was the sum of the individual deficits.

Although NIKL is implemented primarily for defining classes, several definitional features of ABEL concepts could not be represented. Synonymous terms, or terms and their abbreviations, could not be represented as such. In order to represent **ecf** and **extracellular-fluid** as the same entity, we had to specify the latter only as a child of the former; the two classes then "merged." This is a poor representation. Ideally one should be able to access or add structure to either term; this solution only allowed one to hold the structure. Multiple definitions would have also been valuable. For example, the term "acidemia" may be defined as a decreased pH, or as an increase in hydrogen ion concentration. In NIKL there was no method for multiple definitions, nor any approximation. Number intervals would have allowed for more precise specification of low, normal, and high ranges of concentrations. A **normal-serum-k**, for example, could be defined as a **serum-k** with **value** between 3.5 and 4.5. Such intervals were not available for making NIKL definitions and could only be attached to concepts as data. Intervals do seem feasible in a limited form. Sequences of concepts would have also been useful: A therapy may be defined as a set of treatments applied

over time, and an evolving disease may be approximated as a set of symptom states appearing over time. Sequences, too, seem possible to add to NIKL in a limited form. For now, however, NIKL is limited in its definitional capabilities.

Because NIKL is solely for making definitions about concepts and roles, it could not be used to represent the wealth of assertional knowledge in the ABEL knowledge base, including causation, anatomical relations like part-whole and containment, and assertional attributes of all fluids and electrolytes. In order to completely represent all of ABEL's knowledge, there must be in addition to a terminological component like NIKL, a compatible assertional component for both instances and general classes.

4. Conclusion

Although it is difficult to draw definitive conclusions from a single experiment in knowledge representation, our effort to represent the knowledge of the ABEL system in NIKL leads to two clear insights:

1. The lack of a true assertional component in NIKL is devastating to its goals.
2. A richer set of primitive forms of knowledge should be supported by the representation language, even if that compromises the goal of maximum efficiency.

4.1. Terminology and Assertion

We found that, despite our intentions to obey the spirit of the NIKL language, the inability to express assertional information—i.e., information that happens to be true and important about the world, even though it is not part of a definition—tempted and forced us to mis-use NIKL's facilities. Part of the reason for our "moral laxness" in this matter was the observation that the line between terminological and assertional information is much more difficult to draw in practice than is suggested in the literature. Situations arise in which information that appears to be terminological must be represented instead as assertional (and thus, perhaps not represented at all in NIKL), and also in which the terminological component permits the statement of what might really seem as assertional information.

An example of the first case was the definition of acidemia, mentioned above. Consider also the following: It happens to be the case (for deep mathematical reasons that are not clear to the NIKL classifier) that quadrilaterals with pairs of parallel opposite sides are just the same as quadrilaterals with pairs of equal-length opposite sides. The "clean" solution to this—form two distinct concepts in the terminological space and then assert the equivalence of the concepts—leaves terminological knowledge about parallelograms spread between two distinct concepts, with the terminological reasoner unable to pull them together. Instead, one would like some notion of alternative definitions for a concept, where the terminological component would accept as given the equivalence of two definitions even though actually proving that equivalence would be beyond its capabilities. In a workshop of NIKL users two years ago[9], numerous calls for incorporating distinct sets of necessary and/or sufficient conditions for a concept were, we believe, motivated by this same issue.

Assertional information masquerading as terminological arises in the ability of the terminological component to express

the notion of primitive or natural kinds. When we say, for example, that the concept "dog" is a primitive specialization of "animal," it seems awkward to argue that this is purely a matter of terminology. Unlike the non-primitive case wherein a newly-defined term is defined precisely by an existing concept and additional characterizations, defining primitives always involves adding an ineffable characteristic such as "dogness" (in this case). That dogs happen to be animals (or pets, or domesticated animals, or whatever), seems like the essence of a propositional, assertional statement, yet it appears in NIKL as a first-class capability of the terminological component.²

The consequent built-in lack of clarity about just what could be said in the terminological component encourages over-use of the terminological facilities, to squeeze in the domain knowledge in whatever form will fit. Other users of NIKL, for example, the CONSUL project[8], have made similar "out of the spirit" use of NIKL facilities, and have been criticized for it[1]. Even publications in the literature extolling the virtues of NIKL-like representations commit similar sins. In[10], for example, an athlete is defined as a person with an athletic-activity as his hobby. Unfortunately, this is accomplished by giving the concept person a hobby role, which seems hardly part of the definition of the concept person. Nevertheless, until languages like NIKL offer reasonable assertional facilities, well integrated with their terminological components, the temptation for such mis-use will remain overwhelming. Thus far, projects such as KL-TWO[18] promise only a limited solution to this difficulty, because the assertional reasoner in KL-TWO is limited to dealing with discrete individuals and cannot handle quantification.

4.2. The Tradeoff between Expressive Power and Computational Tractability

It is an unfortunate fact of nature that complete inference procedures for languages with even moderate expressiveness are at least exponential[2]. In fact, there appears to be a tradeoff among the competing goals of expressiveness, soundness, completeness and efficiency. The conclusion drawn in the literature—that we must emasculate the expressiveness of languages while retaining soundness and completeness in order to gain control over the inefficiency of their inferences—makes using them for expressing real-world knowledge very difficult.

We believe that there may be other ways to achieve a reasonable tradeoff. Most appealing to us at this time is to relax completeness of the inference system while retaining soundness and (reasonable) efficiency, but increasing expressiveness. This can be accomplished by adding special, limited inference methods to a terminological reasoner that will then support a few specific classes of additional inferences. In encoding the knowledge of ABEL, we would have found the following particularly useful:

1. Declaring and having the classifier make use of the mathematical properties of relations: e.g., transitivity, reflexivity, symmetry. Being able to define one relation as the transitive closure of another would be particularly useful.

²Note also that the very common use of the primitive marker in actual taxonomies will limit the ability of the classifier to determine the proper placement of new descriptions. It is also worth noting that increasing the number of concepts marked primitive decreases the work to be done by the classifier. The common occurrence of primitives may argue, therefore, that the algorithmic efficiency of the classifier is not as important as often suggested.

2. Use of disjoint covers by the classifier, to support reasoning by exclusion.
3. Alternative definitions for a single concept.
4. Intervals of numbers (for expressing numeric ranges).
5. Sequences, especially for representing temporal orders.³

In each of these cases of proposed extensions, it is essential that these special-purpose inferences be made by the terminological component, not a possible assertional one, because we would still insist that the terminological reasoner cannot request the possibly-unbounded services of the assertional in its operation. For example, if we wish to define a D.A.R. as a woman with some ancestor who resided in the United States in 1776, then the system must understand that ancestor is the transitive closure of parent, and this knowledge must be available to the classifier. Because a collection of such special-purpose inference capabilities does not make for a complete inferential system, we would sacrifice completeness for the benefit of gaining some important additional expressive power, while retaining reasonable efficiency. Further efforts are warranted to learn just what such capabilities are critical for representing knowledge in complex domains, and which ones can be efficiently supported by an augmented terminological reasoner. KRYPTON[1], which integrates a powerful resolution-based theorem-prover[16] with its terminological component, may offer enough power in its assertional reasoner, but fails to allow the special-purpose inferences we suggest here as part of its terminological component, thus always forcing the use of general theory resolution, even for relatively simple problems.

Arguments of the form "we tried to use tool *X*, and we didn't do a very good job; therefore tool *X* is flawed" should always be treated with a certain skepticism. Obviously, the flaw is as likely to lie in the user as in the tool. We hope that our supporting arguments provide convincing evidence that in this case, indeed the tools are flawed. In that case, we see compelling evidence for building sophisticated but practical knowledge representation languages at a different point of the expressiveness/tractability curve, as we have suggested.

³This extension is also being investigated in[6].

References

- [1] Ronald J. Brachman, Richard E. Fikes, and Hector J. Levesque. Krypton: A functional approach to knowledge representation. *Computer*, 16(10):67-73, 1983.
- [2] Ronald J. Brachman and Hector J. Levesque. The tractability of subsumption in frame-based description languages. In *Proceedings of the National Conference on Artificial Intelligence*, pages 34-37, American Association for Artificial Intelligence, 1984.
- [3] Ronald J. Brachman and James G. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, 9:171-216, 1985.
- [4] William J. Clancey and Edward H. Shortliffe, editors. *Readings in Medical Artificial Intelligence: The First Decade*. Addison Wesley, Reading, Mass., 1984.
- [5] Ira J. Haimowitz. *Using NIKL in a Large Medical Knowledge Base*. TM 348, Massachusetts Institute of Technology, Laboratory for Computer Science, 545 Technology Square, Cambridge, MA, 02139, January 1988.
- [6] Thomas S. Kaczmarek, Raymond Bates, and Gabriel Robins. Recent developments in NIKL. In *Proceedings of the National Conference on Artificial Intelligence*, pages 978-985, American Association for Artificial Intelligence, 1986.
- [7] Hector J. Levesque. Making believers out of computers. *Artificial Intelligence*, 30(1):81-108, October 1986.
- [8] William S. Mark. Representation and inference in the CONSUL system. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pages 375-381, 1981.
- [9] Johanna D. Moore. NIKL workshop summary, 15-16 July 1986. September 1986. Summary of topics discussed at the meeting among NIKL users and researchers, held at MIT.
- [10] M. G. Moser. *An Overview of NIKL, The New Implementation of KL-ONE*. Technical Report 5421, Bolt, Beranek and Newman, Inc., 1983.
- [11] R. S. Patil, P. Szolovits, and W. B. Schwartz. Information acquisition in diagnosis. In *Proceedings of the National Conference on Artificial Intelligence*, pages 345-348, American Association for Artificial Intelligence, 1982.
- [12] Ramesh S. Patil. *Causal representation of patient illness for electrolyte and acid-base diagnosis*. TR 267, Massachusetts Institute of Technology, Laboratory for Computer Science, 545 Technology Square, Cambridge, MA, 02139, October 1981.
- [13] Ramesh S. Patil, Peter Szolovits, and William B. Schwartz. Causal understanding of patient illness in medical diagnosis. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pages 893-899, 1981.
- [14] W. B. Schwartz, R. S. Patil, and P. Szolovits. Artificial intelligence in medicine: where do we stand. *New England Journal of Medicine*, 316:685-688, 1987.
- [15] Brian C. Smith. *Reflection and Semantics in a Procedural Language*. TR MIT/LCS/TR-272, Massachusetts Institute of Technology, Laboratory for Computer Science, 545 Technology Square, Cambridge, MA, 02139, 1982.
- [16] M. E. Stickel. A nonclausal connection-graph resolution theorem-proving program. In *Proceedings of the National Conference on Artificial Intelligence*, pages 229-233, 1982.
- [17] P. Szolovits, R. S. Patil, and W. B. Schwartz. Artificial intelligence in medical diagnosis. *Annals of Internal Medicine*, 108:80-87, 1988.
- [18] Marc B. Vilain. The restricted language architecture of a hybrid representation system. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 547-551, 1985.
- [19] William A. Woods. What's in a link: Foundations for semantic networks. In Bobrow and Collins, editors, *Representation and Understanding*, pages 35-82, Academic Press, New York, 1975.